

基于异配特征融合的图神经网络

黄治玮¹

HUANG Zhiwei

摘要

在异配性环境下, 节点更倾向于与不同标签的节点相连, 导致图神经网络往往表现出次优性能。当前关于图异配性的研究主要关注于聚合校准或邻居扩展, 并通过利用节点特征或结构信息来解决异配性问题, 从而改进 GNN 表示。但异配性数据同时也包含着大量的特性信息, 基于此, 文章提出了一种融合异配特征的图神经网络——FusionH-GNN, 通过构造一个全新的图结构, 利用邻居的分布整合异配特征, 以增强共享相似语义特征的节点之间的连通性。并在同配和异配基准数据集上进行了节点分类任务的实证评估, 将 FusionH-GNN 与经典 GNN 基准方法进行比较, 证实了其在改进图表示方面的有效性。

关键词

图神经网络; 同配性; 模式识别; 图表示学习

doi: 10.3969/j.issn.1672-9528.2025.12.015

0 引言

图结构数据广泛存在于社交网络^[1]、推荐系统^[2]等各类应用中, 这类数据中的实体通过特定关系相互连接。为从这些关系图中高效学习数据表示, 图神经网络^[3-5]应运而生并取得了显著进展。GNNs 在图结构数据上的有效性在很大程度上依赖于同配性假设, 该假设认为相连的节点倾向于相似^[6-7]。在这一假设下, 信息从具有相似特征的邻居节点中聚合, 促进了局部结构内有效的信息交换。然而, 在许多现实世界场景中, 同配性假设并不总是成立, 异配节点之间存在连接, 这种情况被称为异配性或非同配性。异配性对 GNNs 的性能构成了限制, 因为它涉及从异配节点聚合信息, 在聚合过程中引入了额外的噪声。

为解决图中的异配性问题, 已有多种方法被提出。部分方法通过整合高阶邻居并赋予不同权重^[8-10]、将邻接矩阵作为节点的新特征^[11]、或传播先验信念估计^[12], 以校准聚合过程。尽管这些聚合校准方法在异配图取得了一定改进, 但它们主要聚焦于局部邻居, 难以有效聚合全局邻居的信息。因此, 另一些方法通过引入新型滤波器^[13]或构建新的图结构^[14], 从全局邻居中提取并聚合信号。虽然在许多基于图的机器学习任务中, 异配性常被视为消息聚合过程中的噪声^[14], 但在某些场景下, 异配性的图结构可能蕴含重要的语义信息。例如在学术合作网络中, 来自不同学科的学者通过学术合作相连, 网络中的异配性反映了跨学科合作的意义, 比如生物学以及计算机信息技术。在这种情况下, 高度的异

配性表明不同领域之间存在丰富的思想与方法交流, 能够促进创新并拓展研究视角。尽管 GNNs 的性能常因异配性而受限, 但异配性图中蕴含的语义信息不容忽视。

针对以上挑战, 本文提出了一种融合异配性特征的图神经网络模型——FusionH-GNN。该模型从异配性中提取特征信息, 具体描述了节点邻居属于特定类别的概率分布。这一信息不仅为同质性或异配性程度的分析提供了依据, 还深化了对节点异配性上下文的理解, 从而提升模型的表达能力。

1 预备知识

1.1 图结构数据

本文将 $G = (V, E)$ 记为一个无向图 G , 其中节点集为 V ; 边集为 E 。设 A 为邻接矩阵, 其中 $A_{ij}=1$ 或 $A_{ij}=0$ 表示连接节点 v_i 和 v_j 的边 e_{ij} 的存在或不存在。

对于每个节点 v_i , 其 1 跳邻居集定义为 $N_i = \{v_j | A_{ij}=1\}$ 表示, 并用 d_i 表示其节点度。并将节点特征表示为 $X = \{x_1, x_2, \dots, x_n\} \in \mathbb{R}^{N \times M}$; 标签表示为 $Y = \{y_1, y_2, \dots, y_n\} \in \mathbb{R}^{N \times 1}$, 其中 N 和 M 分别是 G 中的节点数和特征数。

1.2 图神经网络

图神经网络是专门用于处理图结构数据的神经网络, 基于消息传递机制 GNNs 使节点能够与其邻居交换信息。在 GNN 的第 l 层中, 节点 v_i 的 z_i^l 更新为:

$$z_i^l = \text{UPDATE}^l(z_i^{l-1}, \text{AGGREGATE}^l(z_i^{l-1}, \{z_j^{l-1} | v_j \in N_i\})) \quad (1)$$

式中: AGGREGATE($\cdot$) 描述了从节点 v_i 及其邻居 N_i 聚合信息的过程; UPDATE($\cdot$) 描述了 z_i^l 如何基于自身表示 z_i^{l-1} 及聚合的邻居表示更新。

1. 贵州财经大学信息学院 贵州贵阳 550025

图卷积网络是 GNNs 的代表之一，其第 l 层的输出可以表示为 $Z^l = \sigma(\tilde{A}Z^{l-1}W^{l-1})$ ，激活函数 (σ) 被应用于归一化邻接矩阵 $\tilde{A} = D^{-\frac{1}{2}}\tilde{A}D^{-\frac{1}{2}}$ ，其中 $\tilde{D} = D + I_n$ 和 $\tilde{A} = A + I_n$ 分别表示带自环的度矩阵和邻接矩阵。

1.3 异配性特征

在异配性图中，不同标签的节点更倾向于相互连接。因此，众多研究将异配性视为图表示学习的核心挑战，并提出各类方法，通过挖掘节点特征或结构信息来应对这一问题。然而，对学术合作网络的分析表明，异配性中本可助力图表示学习的关键信息，尚未得到充分利用。基于此，本文首先给出异配特征的明确定义，进而深入探讨其在改进图学习效果中的具体益处

对于图 G 中的任意节点 u ，将异配特征 $\mathcal{H}_u$ 定义为其邻居分布，描述节点呈现异配性的方式：

$$\mathcal{H}_u = [p_1, p_2, \dots, p_c], \text{ 其中 } p_i = \frac{|\{v \in \mathcal{N}_u, y_v = i\}|}{|\mathcal{N}_u|} \quad (2)$$

式中： $\mathcal{H}_u$ 中的第 i 个元素 p_i 描述了节点邻居 $\mathcal{N}_u$ 属于类别 i 的概率。与节点同配性的定义相比，异配信息更能表征节点的异配性上下文，对于任意具有类别 i 的节点 u ，当其同配性非常高（如 $h \rightarrow 1$ ）时， $\mathcal{H}_u$ 近似于全零向量，且 $p_i \rightarrow 1$ 。相反，低同配性导致 $\mathcal{H}_u$ 中的 $p_i \rightarrow 0$ 。在这种情况下， $\mathcal{H}_u$ 的其他条目揭示了节点呈现异配性的方式，而 h 作为一个单一标量值，难以全面表征节点的异配性上下文。

接下来，为了利用异配特征，通过连接高相似度 $\mathcal{H}$ 的节点，创建新的邻接矩阵 A' ，其中每个元素 A'_{ij} 可以表示为：

$$A'_{ij} = \mathbb{1}(\cos(\mathcal{H}_i, \mathcal{H}_j) > \delta) \quad (3)$$

式中： $\mathbb{1}$ 是一个指示函数，当节点 i 和 j 的异配信息的余弦相似度大于阈值 δ 时取值为 1，否则取值为 0。

2 FusionH-GNN 模型

在本节中，提出异配特征融合图神经网络（fusion heterophilous features graph neural networks, FusionH-GNN），该方法基于之前对异配的分析。

利用图的异配性构造新的图结构，其包含了异配特征信息。图 1 展示了通过整合异配信息 $\mathcal{H}$ 构造新邻接矩阵 A' 的过程。

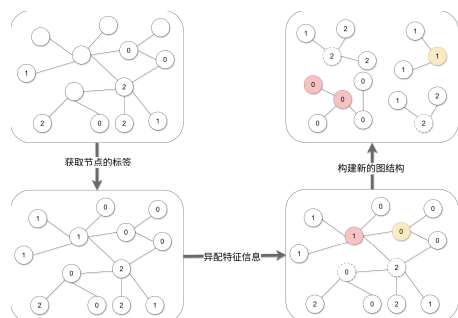


图 1 利用已知节点构造异配特征

首先通过一个基础模型来获取所有标签 Y' 。其次遵循传统的图上半监督学习的设置，在训练数据 X_{train} 和 Y_{train} 上训练现成模型来估计所有节点的标签 A' 。

$$\min_M \mathcal{L}(M(X_{\text{train}}), Y_{\text{train}}), Y' = M(X) \quad (4)$$

式中： $M(\cdot)$ 表示一个现成模型； $\mathcal{L}$ 表示损失函数。获取所有估计标签 Y' 后，通过使用式 (4)，根据邻居分布推导出所有节点的异配信息 $\mathcal{H}$ 。最后，基于 $\mathcal{H}$ ，通过使用式 (3) 保留高相似度的连接来构造 A' 。

对于表现出高同配性或异配性的节点，分别在图 1 中标记为深色节点或浅色节点，利用基础的模型能有效将它们分组到同一类别中。

此外，本文方法还可以连接具有半同配性的节点，在图 1 中用虚线表示。这些中等同配性节点作为跨越两个类别的边界节点，共享相似的语义。虽然将边界节点与类别边界内的节点连接会增加同配性，但可能对图学习产生不利影响。这是因为边界节点与其他类型节点连接的影响，导致出现不同的模式表现。新邻接矩阵在边界节点之间建立连接，并促进相似语义节点间的消息传递，从而在不同的同配性水平上增强图学习的有效性。尽管这种标签估计并非完美，可能导致异配信息 $\mathcal{H}$ 的偏差估计，但 $\mathcal{H}$ 本身具有容错性，能够适应相似节点之间邻居分布的差异。

然后针对新旧两种图结构进行通道融合。虽然新构造的邻接矩阵 A' 包含丰富的异配信息并提高了同配性程度，但原始邻接矩阵 A 同样具有特征信息。为保留原始图拓扑信息作为补充，采用晚融合策略对这两种图拓扑结构进行处理。具体而言，对于第 $l-1$ 层的节点嵌入 Z_{l-1} ，在 A' 和 A 上执行图卷积，分别获取下一层更新后的节点嵌入 Z'_{new} 和 Z'_{old} ：

$$Z'_{\text{new}} = g(\hat{A}Z^{l-1}W^{l-1}), Z'_{\text{old}} = g(\tilde{A}Z^{l-1}W^{l-1}) \quad (5)$$

式中： $\hat{A}$ 和 $\tilde{A}$ 分别表示 A' 和 A 的归一化邻接矩阵； $g(\cdot)$ 表示激活函数； W^{l-1} 表示权重矩阵。

由于异配特征信息是新添加的内容，其对最终结果的贡献可能因具体数据集而异，引入平衡参数 λ 对来自两个通道的嵌入进行晚融合，可表示为：

$$Z^l = \lambda Z'_{\text{new}} + Z'_{\text{old}} \quad (6)$$

平衡参数 λ 用于调节学习过程中纳入异配信息的比例。较高的 λ 强调异配信息，而较低的 λ 强调原始拓扑信息。选择晚融合而非早期组合两种邻接矩阵，是有意保持异配信息和拓扑信息彼此不受干扰。这种方法确保每个部分的独特贡献能够被单独捕获，从而使每个部分充分表达其信息通道的细微差别。

3 实验设置

3.1 基准方法

将 FusionH-GNN 与多种方法进行基准测试, 包括一般 GNNs 以及异配性为导向的 GNNs。具体而言, 使用 GCN^[3]、GAT^[4]、GraphSage^[5] 作为一般 GNNs 的基准方法。对于异配性导向的 GNNs, 使用 Mixhop^[8]、GloGNN^[15]。

3.2 数据集

在 3 个同配性数据集和 3 个异配性图数据集上测试所有方法。数据集的统计数据如表 1 所示, 其中同配性图 (如 Cora、CiteSeer 和 PubMed) 表现出显著较高的同配性, 相比之下异配性图 (Actor、Texas 和 PubMed) 较低。对每个方法运行 10 次实验, 遵循公共随机划分比例 60%、20%、20% 用于训练、验证和测试集, 取验证集上最优轮次所对应的测试集表现作为一次实验的代表, 最后统计 10 次实验结果的均值和标准差。

表 1 数据集信息

Dataset	Cora	CiteSeer	PubMed
#Nodes	2 708	3 327	19 717
#Edges	10 556	9 104	88 648
#Features	1 433	3 703	500
#Classes	7	6	3
h_{edge}	0.81	0.74	0.8
h_{node}	0.83	0.72	0.79

Dataset	Actor	Texas	Wisconsin
#Nodes	7 600	183	251
#Edges	37 526	492	750
#Features	932	1 703	1 703
h_{edge}	0.22	0.06	0.17
h_{node}	0.22	0.10	0.15

3.3 实现方式

为生成邻居分布的标签预测, 选择 ACM-GCN 作为现成模型。鉴于方法的灵活性 (可集成到任何 GNN 方法中), 采用 GCN 作为主干模型, 并通过高通滤波器进行增强。然后, 在 6 个数据集上运行 FusionH-GNN 及所有基准方法, 使用 PyTorch 在 Nvidia Tesla P100 上实现。模型训练至收敛, 采用早停耐心值为 40 个周期。

对于所有方法, 均使用超参数搜索最佳超参数, 包括隐藏层的嵌入大小从 {64, 128, 256}、学习率为 $\{10^{-2}, 10^{-3}\}$ 、权重衰减为 $\{5 \times 10^{-4}, 5 \times 10^{-5}\}$ 、层数为 {1, 2} 以及丢弃率为 {0.3, 0.4, 0.5, 0.6, 0.7}, 这些超参数在每个数据集的验证集上调整。对于 FusionH-GNN, 特别搜索阈值 δ 为 {0.5, 0.8, 0.9, 1.0} 以及 λ 为 {1, 0.5, 10^{-1} , 5×10^{-2} , 1×10^{-2} , 1×10^{-3} , 1×10^{-4} }, 以控制涉

及到的异配特征的信息量。

3.4 性能

表 2 展示了 FusionH-GNN 与其他基准的比较, 所有方法均通过节点分类任务进行评估。每个方法的平均排名基于 6 个数据集计算。对于每个数据集, 最佳表现方法用粗体标出, 次优方法用下划线标出。对于通过异配特征信息增强的方法记为 -HF。

首先, 提出的 FusionH-GNN 总体上优于其他方法, 凸显了将异配特征信息融入图学习的有效性。FusionH-GNN 在同配数据集和异配数据集上的性能均得到了提升, 说明通过异配特征信息构造的新矩阵能够在不同的同配性水平上提升特征提取的效果。其次, 异配信息的效果还体现在一般 GNN 基准的性能提升上, 大多数数据集上的性能均有所改善。最后, 与一般 GNN 基准如 GCN 和 GAT 相比, 异配图神经网络在同配性图上表现较差, 但在异配性图上仍表现出较强性能, 对于在消息传递中聚合 k 跳邻居的模型 Mixhop, 尽管其性能相较 GCN 有所提升, 但与利用全局相似节点的 GloGNN 相比仍存在较大差距。这表明, 对于异配图, 仅关注 k 跳邻居的聚合不如构建新图结构以捕捉语义相似的邻居有效。

表 2 不同异配性数据集上的表现

Methods	Cora	CiteSeer	PubMed
GCN	<u>85.3 ± 0.9</u>	76.4 ± 0.8	86.1 ± 0.4
GCN-FH	86.5 ± 0.5	77.3 ± 0.6	89.0 ± 0.3
GAT	84.3 ± 1.5	77.0 ± 1.1	85.3 ± 0.4
GAT-FH	84.8 ± 0.4	76.0 ± 0.4	<u>86.7 ± 0.3</u>
GraphSage	84.1 ± 1.4	80.1 ± 5.1	79.3 ± 0.3
GraphSage-FH	83.8 ± 0.6	83.8 ± 0.5	80.8 ± 0.3
Mixhop	75.3 ± 0.2	87.5 ± 0.5	77.6 ± 0.2
Mixhop-FH	73.3 ± 0.5	<u>84.2 ± 0.4</u>	78.6 ± 0.3
GloGNN	76.4 ± 0.2	74.6 ± 0.3	75.5 ± 0.7
GloGNN-FH	76.8 ± 0.5	76.8 ± 0.5	76.8 ± 0.5

Methods	Actor	Texas	Wisconsin
GCN	71.2 ± 0.6	75.7 ± 2.1	67.1 ± 1.9
GCN-FH	72.5 ± 0.5	78.1 ± 1.9	69.2 ± 1.8
GAT	72.0 ± 0.6	77.3 ± 1.8	69.8 ± 1.5
GAT-FH	75.8 ± 0.4	79.2 ± 1.4	70.5 ± 1.3
GraphSage	70.4 ± 0.7	74.2 ± 2.0	71.5 ± 1.8
GraphSage-FH	72.1 ± 0.6	76.3 ± 1.5	72.7 ± 1.6
Mixhop	80.3 ± 0.4	79.7 ± 1.2	83.1 ± 1.0
Mixhop-FH	<u>82.5 ± 0.4</u>	81.2 ± 1.2	<u>83.9 ± 1.0</u>
GloGNN	82.4 ± 0.2	<u>82.6 ± 0.3</u>	82.5 ± 0.7
GloGNN-FH	85.8 ± 0.5	84.8 ± 0.5	84.8 ± 0.5

4 结论

本文提出了一种创新的图神经网络方法——FusionH-GNN,旨在充分利用图中异配性所蕴含的语义信息。与以往将异配性视为干扰或噪声不同,本文将其视作潜在的、有价值的结构特征,通过分析节点邻居的标签分布,挖掘节点间更深层次的语义联系。FusionH-GNN通过构建新的图结构,有效整合了异配特征,增强了具有相似语义属性节点之间的连通性,从而提升了图表示的质量和下游任务的性能。

在多个同质性与异配性基准数据集上的节点分类实验验证了本方法的有效性。与经典的GNN模型相比,FusionH-GNN在不同类型的图结构中均展现出更优的泛化能力和鲁棒性,尤其在异配性较强的图中优势更为明显,表明其在异构环境下的图学习任务中具有良好的适应性和实用价值。

FusionH-GNN在实际应用中仍存在一定的局限性。例如,该方法在图结构构建过程中依赖于对节点标签分布的预估,这在部分无标签或标签不完全的数据集中可能影响其性能。因此,未来的研究可进一步探索如何在弱监督或无监督的条件下更高效地建模异配性信息,以提升方法的适用范围与自动化程度。

参考文献:

- [1]GUO Z W, WANG H. A deep graph neural network-based mechanism for social recommendations[J]. IEEE transactions on industrial informatics, 2021, 17(4): 2776-2783.
- [2]WU S, TANG Y Y, ZHU Y Q, et al. Session-based recommendation with graph neural networks[C]//Proceedings of the 33rd AAAI Conference on Artificial Intelligence and 31st Innovative Applications of Artificial Intelligence Conference and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence. Palo Alto: AAAI press, 2019: 346-353.
- [3]KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks[EB/OL]. (2017-02-22)[2025-12-02].<https://doi.org/10.48550/arXiv.1609.02907>.
- [4]VELIČKOVIĆ P, CUCURULL G, CASANOVA A, et al. Graph attention networks[EB/OL]. (2018-02-04)[2025-12-02].<https://doi.org/10.48550/arXiv.1710.10903>.
- [5]HAMILTON W L, YING R, LESKOVEC J. Inductive representation learning on large graphs[C]//NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems.Brussels:ACL,2017:1025-1035.
- [6]ZHAO G P, WANG T, LI Y D, et al. Neighborhood pattern is crucial for graph convolutional networks performing node classification[J]. IEEE transactions on neural networks and learning systems, 2024, 35(6): 8456-8469.
- [7]ZHENG X, WANG Y, LIU Y X, et al. Graph neural networks for graphs with heterophily: a survey[EB/OL]. (2024-02-25)[2025-12-02].<https://doi.org/10.48550/arXiv.2202.07082>.
- [8]ABU-EL-HAJJA S, PEROZZI B, KAPOOR A, et al. MixHop: higher-order graph convolutional architectures via sparsified neighborhood mixing[EB/OL].(2019-06-19)[2025-12-02].<https://doi.org/10.48550/arXiv.1905.00067>.
- [9]ZHU J, YAN Y J, ZHAO L X, et al. Beyond homophily in graph neural networks: current limitations and effective designs[C]//NIPS'20: Proceedings of the 34th International Conference on Neural Information Processing Systems. NewYork: ACM, 2020:7793 - 7804.
- [10]WANG X, BO D Y, ZHU M Q, et al. AM-GCN: adaptive multi-channel graph convolutional networks[C]//KDD '20: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. NewYork:ACM,2020: 1243-1253.
- [11]LIM D, HOHNE F, LI X Y, et al. Large scale learning on non-homophilous graphs: new benchmarks and strong simple methods[C]//NIPS'21: Proceedings of the 35th International Conference on Neural Information Processing Systems. NewYork: ACM, 2021: 20887-20902.
- [12]ZHU J, ROSSI R A, RAO A, et al. Graph neural networks with heterophily[EB/OL]. (2021-06-14)[2025-12-02].<https://doi.org/10.48550/arXiv.2009.13566>.
- [13]JAMADANDI A, RUBIO-MADRIGAL C, BURKHOLZ R. Spectral graph pruning against over-squashing and over-smoothing[EB/OL].(2024-10-30)[2025-12-02].<https://doi.org/10.48550/arXiv.2404.04612>.
- [14]MAO H T, CHEN Z K, JIN W, et al. Demystifying structural disparity in graph neural networks: can one size fit all?[C]//NIPS'23: Proceedings of the 37th International Conference on Neural Information Processing Systems. NewYork:ACM,2023: 37013-37067.
- [15]LI X, ZHU R Y, CHENG Y, et al. Finding global homophily in graph neural networks when meeting heterophily[EB/OL].(2022-05-15)[2025-12-02].<https://doi.org/10.48550/arXiv.2205.07308>.

【作者简介】

黄治玮(2000—),男,贵州思南人,硕士研究生,研究方向:计算机应用技术。

(收稿日期:2025-07-03 修回日期:2025-12-15)